JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 1993, 60, 469-472 NUMBER 2 (SEPTEMBER)

INTERNAL STATES: NECESSARY BUT NOT SUFFICIENT

ARMANDO MACHADO

DUKE UNIVERSITY

Staddon's article is timely, refreshing, and provocative; timely because it invites us to rethink our basic assumptions, always a healthy activity in periods of crisis; refreshing because it suggests new avenues for behavior analysis, such as the history-based concept of internal state; and provocative because it questions a cherished discrimination made by many behavior analysts, that between organism- and environment-based accounts of behavior. That much is positive. However, as I show below, Staddon misinterpreted Darwin, misconstrued organism-based accounts, and did not identify the major source of the current problems in behavior analysis.

On Darwin and Darwinism. Coming from an author who has written extensively on the relations between evolutionary biology and psychology (e.g., Staddon & Simmelhag, 1971), it is surprising to read that "Darwin dealt primarily with environmental determinants only because behavior was not his primary object of study.... When Darwin did deal with behavior, his theorizing was more 'organism based' than might be suspected" (p. 439). The statement is inaccurate on two grounds. First, no reader of the *Origin* can fail to recognize in Darwin's account of the origin of species and their various adaptations, how delicate, rich, complex, and fundamental was the role of the environment. As a distinguished historian of biology put it, "Almost his greatest service to biology was that he made biologists realize as they never did before the vast importance of the environment" (Russell, 1917/1982, pp. 232–233). Parenthetically, if we replace "biology" with "psychology," exactly the same can be said about B. F. Skinner.

Second, Darwin suggested one and the same process to explain the evolution of both the "corporeal structures" and the behavior of organisms. The process is eloquently summarized in his concluding comments in the chapter on instinct:

> it may not be a logical deduction, but to my imagination it is far more satisfactory to look at such instincts ... as small consequences of one general law, leading to the advancement of all organic beings, namely, multiply, vary, let the strongest live and the weakest die. (Darwin, 1859/1987, p. 263).

That behavior is but another "organ" whereby animals adapt to their environments was also frequently echoed by the modern founders of ethology, Konrad Lorenz and Nikolaas Tinbergen.

Staddon's misinterpretation of Darwin highlights a more general issue: Darwinism is not a single theory but a set of theories (e.g., Mayr, 1988, chap. 4), and when different authors emphasize different theories they are likely to reach opposite conclusions. For example, Darwin's account of vestigial characters, or the phases of embryological development, appeals to common descent; in the procrustean framework of organism- versus environment-based accounts, we would probably classify common descent as an organism-based theory. Had I given an example of the workings of natural selection, such as the cell-making instinct of the hive-bee (Darwin, 1859/1987, pp. 247–262), we would certainly lean towards the environment pole. Hence, let Darwin himself clarify what he considered of utmost importance:

> It is generally acknowledged that all organic beings have been formed on two great laws—Unity of Type, and the Conditions of Existence.... On my theory, unity of type is explained by unity of descent. The expression of conditions of existence ... is fully embraced by the principle of natural selection.... Hence, in fact, *the law of the Conditions of Existence is the higher law; as it includes, through the inheritance of former variations and adaptations, that of Unity*

I thank Nancy Innis, Jennifer Higa-King, Mark Cleveland, and John Staddon for many discussions of these issues. Correspondence concerning this article may be sent to Armando Machado, Department of Psychology: Experimental, Duke University, Durham, North Carolina (E-mail: Armando@Psych.Duke.Edu).

of Type. (Darwin, 1859/1987, p. 233, italics added)

Organism- versus environment-based explanations. Consistent with his approach, Staddon avoids explicit definition of his terms. However, he often writes as if subject matter or terminology alone were sufficient criteria to classify a behavioral account as organism or environment based. Thus, when commenting on the behavior of Darwin's ants, Staddon classifies as organism based an account that (a) defines two competing action patterns and (b) suggests a hypothesis concerning the effects of various manipulations on the relative strength of the behaviors involved. Granted that (a) and (b) are not prerogatives of any type of account, Staddon must have relied on subject matter, instinctive behavior, to classify the account. On the other hand, when discussing the cumulative trace model, Staddon argues that a model that assumes an exponential decay of the effects of reinforcement qualifies as organism based because such a model is in fact a memory trace model. As no additional property of memory (e.g., storing, retrieving) or its trace was presumed, I conclude that the shift from an environment- to an organism-based model was achieved through relabeling.

Neither subject matter nor terminology are appropriate criteria, however. Instinctive behavior is no more typical of an organism than is noninstinctive behavior, and relabeling alone cannot change the nature of an explanation. What, then, is an organism-based account? It is an account based on inferred mental structures and ad hoc internal actions that are metaphorically extended from the outside world to the domain of the mind. To illustrate, consider the following account of a phenomenon well known to Staddon—transitive "inference" in pigeons (e.g., Fersen, Wynne, Delius, & Staddon, 1991). Present Stimuli A and B to a pigeon and reinforce only pecks on A (A+B−); when the discrimination is learned, introduce a new pair, B and C, and reinforce responses on B (B+C−). Repeat the procedure two more times by presenting C+D− and D+E−. At the end of training, when confronted with the new pair BD, pigeons frequently choose B, thus showing transitive "inference" (Fersen et al., 1991). One explanation of this result is the mental line theory. During training, the bird orders the various stimuli along a mental line: A is anchored at the left end, B to its right, C and D next, and finally E at the right end of the line. Given a choice between B and D the bird scans the line from left to right and chooses the leftmost stimulus. A novel prediction follows: Reaction time should be an increasing function of the linear (mental) distance between the stimuli facing the subject.

The mental line theory qualifies as organism based not because of its subject matter or its labels but because the structures involved (the mental line) and the presumed actions (ordering the stimuli along the line, scanning the line) are not reducible to the history of the bird. Also, the aforementioned prediction concerning reaction time is derived from properties of the representation, not from historical variables.

The nature of a behavioral theory generally mirrors the practices and purposes of its respective theorist. Trying to characterize what an animal has learned, organism-based theorists vary the environmental context and observe the resulting behavioral changes. Because stimuli and responses are not defined functionally, the burden of explanation is placed in internal constructs and structures; knowledge acquisition in the form of a new representation constitutes the invariant of the learning episode. On the other hand, environment-based theories emphasize under what conditions learning takes place; hence, these theorists manipulate histories and record the ensuing behavioral changes. Internal processes are avoided by redefining stimulus and response classes, the invariants of the learning episode (Zuriff, 1985).

Organism-based accounts may be compared with Ptolemy's geocentric model of the movement of celestial bodies. Epicycles, eccentrics, equants, and the like were contrived by Ptolemy and his predecessors to account for observational data and to predict eclipses and other astronomic events. However, none of the geometric devices was derived from more basic, physical laws; that had to wait for Newton. Analogously, most cognitive, organism-based constructs are developed to describe and predict behavioral data but are not rooted in more basic principles, whether neurophysiological

or phylogenetic/ontogenetic in character. Moreover, geometry, the revered language of ancient times, has its modern equivalent in the revered language of information processing.

Like Ptolemy's *Almagest,* which remained *the* astronomy book for almost 1,400 years because it predicted the movement of the planets with amazing accuracy, organism-based theories will endure until alternative, history-based accounts with equivalent or superior predictive power are advanced. In the meantime, the interpretations offered by behavior analysts, who, metaphorically speaking, have waved Newton but not delivered the *Principia,* remain prone to the criticism that they are just-so stories (Dennett, 1983), or that they leave their subjects "lost in history" (Roitblat, 1982, p. 394).

Behavior analysis, the source of its problems and its future. Although Staddon did not appropriately characterize organism-based explanations, he nevertheless prescribed part of the right therapy to relieve behavior analysis from its current malaise,[1] namely, the history-based concept of internal state (see also Leyton, 1992, chap. 4, and particularly Casti, 1989, chap. 3). If recent events, such as *JEAB*'s special issue on behavioral dynamics, are reliable predictors of the future, then the language of internal states—always defined as sets of equivalent histories—will become part of the vocabulary of behavior analysis.

Will internal states and corresponding dynamic analyses suffice to rescue the field? I doubt it, because internal states are the right medicine, but only for part of the problem. As I see it (Machado, 1992), the major difficulties facing behavior analysis do not come from what behavior analysts do, but from what they do not do that is worth doing. In dealing almost exclusively with reinforcement learning in pigeons, rats, and humans, the field has ignored many types of behavior that do not fit the framework of operant behavior (e.g., fixed action patterns, song learning in birds, spatial learning and memory) and many animal species (e.g., most invertebrates) in which learning plays only a minor role in their adaptation to the environment. Hence, the conceptualization of behavior is as narrow as its field of inquiry.

Ultimately, behavior analysis must choose between two courses of action: (a) In the spirit of Branch's (1992) editorial, behavior analysis can *generalize* its field of inquiry in order to include other types of behavior and other animals species, or (b) it can *specialize* in the study of reinforcement learning. The latter alternative is more likely than the former, however, for the following reason. Scientists organize themselves in small and local research groups. This demic structure of science (Hull, 1988) predicts that in times of intense bilateral criticism, when research domains rarely overlap and cross-references in scientific communication are not abundant, speciation is likely to occur.

If speciation does occur, then behavior analysis faces some good and some bad news: The bad news is that most small populations simply die out without leaving a trace, except as fossils in the record of history; the good news is that most evolutionary advances do occur in small and cohesive populations.

[1] Other authors have used different metaphors to characterize the current status of behaviorism. For example, Catania (1987) refers to it as a (necessarily temporary) eclipse; Amsel (1989), using a parliamentary metaphor, describes it as the "loyal opposition."

REFERENCES

Amsel, A. (1989). *Behaviorism, neobehaviorism, and cognitivism in learning theory: Historical and contemporary perspectives.* Hillsdale, NJ: Erlbaum.

Branch, M. N. (1992). On being narrowly broad. *Journal of the Experimental Analysis of Behavior,* **57,** 1–4.

Casti, J. L. (1989). *Alternate realities: Mathematical models of nature and man.* New York: Wiley.

Catania, A. C. (1987). Some Darwinian lessons for behavior analysis: A review of Bowler's *The Eclipse of Darwinism. Journal of the Experimental Analysis of Behavior,* **47,** 249–257.

Darwin, C. (1987). *The origin of species.* New York: Penguin Books. (Original work published 1859)

Dennett, D. C. (1983). Intentional systems in cognitive ethology: The "Panglossian paradigm" defended. *Behavioral and Brain Sciences,* **6,** 343–390.

Fersen, L., Wynne, C. D. L., Delius, J. D., & Staddon, J. E. R. (1991). Transitive inference formation in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes,* **17,** 334–341.

Hull, D. L. (1988). *Science as a process: An evolutionary account of the social and conceptual development of science.* Chicago: University of Chicago Press.

Leyton, M. (1992). *Symmetry, causality, mind.* Cambridge, MA: MIT Press.

Machado, A. (1992). *The future of behaviorism: An insider's perspective.* Paper presented at the 25th International Congress of Psychology, Brussels, Belgium.

Mayr, E. (1988). *Toward a new philosophy of biology: Observations of an evolutionist.* Cambridge, MA: Belknap.

Roitblat, H. L. (1982). The meaning of representation in animal memory. *Behavioral and Brain Sciences,* **5,** 353–406. (Includes commentary)

Russell, E. S. (1982). *Form and function: A contribution to the history of animal morphology.* Chicago: University of Chicago Press. (Original work published 1917)

Staddon, J. E. R., & Simmelhag, V. L. (1971). The "superstition" experiment: A reexamination of its implications for the principles of adaptive behavior. *Psychological Review,* **78,** 3–43.

Zuriff, G. E. (1985). *Behaviorism: A conceptual reconstruction.* New York: Columbia University Press.